Matthias Standfest

Research Assistant

Telephone:+41 44 633 70 69
Email: standfest@arch.ethz.ch

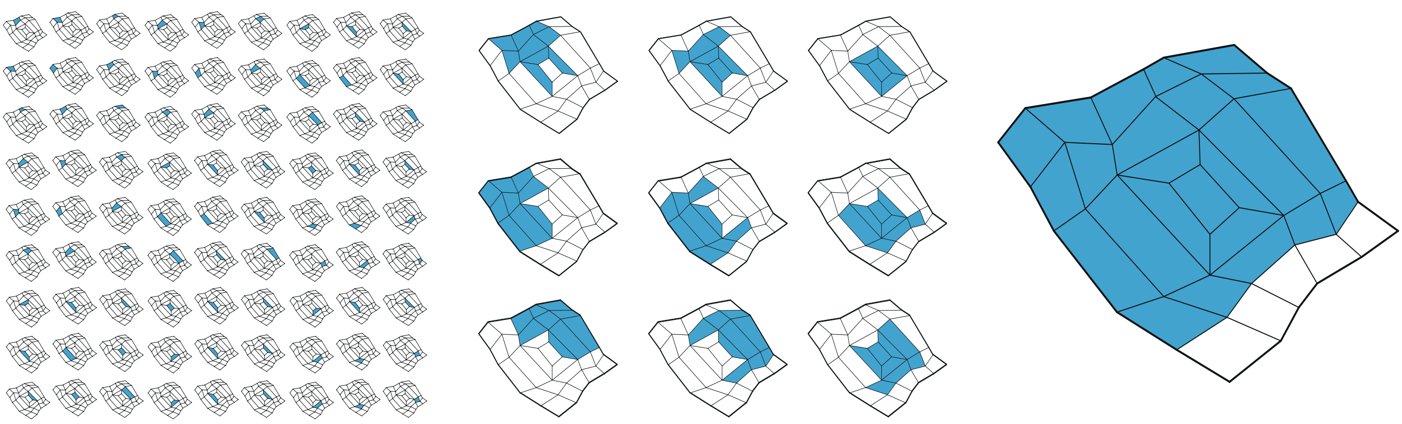
Chair of Information Architecture

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Data Mining Meaningful Clusters of Geometry; Applying Probabilistic Concepts of Architectural Form

The work investigates the problem of how to interpret the meaning of urban geometry, thus the architectonic form. The research is focusing on how to use modern statistical methods known from Machine Learning to compare unlabeled geometric data. Therefore a symmetric encoding from quad mesh representation into feature vectors for statistical processing is developed. This data is further used to train Deep Learning networks to detect frequent patterns (emergent features) in geometric configurations. A case study of 15.000 apartments is used to compare the performance of this contextual clustering approach in comparison to intrinsic feature parameters. This thesis should determine if a correlation between purely mathematical geometry and the qualitative perception of architectonic form can be statistically reproduced.

One intrinsic problem of the architectural design process is the issue of how to compare different designs with each other. As a necessity for any synthesis driven design approach this wicked problem has yet to be fully automatized, especially in consideration that existing methods fail in practicability because of expensive manual data pre-processing. So far those methods can be divided into three main groups corresponding to the particular kinds of compared data: shape determining rules instead of a shape itself; key figures of arbitrary features as a substitute for an underlying geometry; and polygon meshes including certain topological information. The trend is clearly towards minimizing the costs for useful **comparisons by minimizing the amount of semantic information needed, which reflects the algorithmic modelling efforts in state of the art data analysis.** To contribute to this development, the conducted research is focusing on the problem of comparing unlabelled plain geometry data by using both many layered neural networks and n-gram inspired mesh encoding. This topology oriented approach in data analysis extends hierarchical analysis of shape structures as well as probabilistic models for component based shape synthesis. Therefore, derived from well-investigated deep learning architectures, a custom many layered **Convolutional Neural Network** is developed. In addition to that, a mesh sub-clustering without the need for semantically correct shape segmentation will ensure a strictly unsupervised workflow. In order to measure the two most relevant parameters - prediction quality and clustering quality – multiple experiments on a data sets of 15.000 3D HQ apartment geometries are going to be conducted. These different scenarios will help to determine the best performing data labels (e.g. UAVs) and the optimal process parameters. Finally an analysis of the created CNN kernels will show statistical relevant geometric clusters in urban design. I expect the series of experimental research to produce necessary findings required for the future development of unsupervised mesh indexing, led by an estimation of appropriate data volume and variety. While the prediction quality should be at least at five sigma, the mapping will show from noise clearly distinguishable clustering of sub meshes. Furthermore the **computing costs are going to show advantages in efficiency over semi-automatic methods or online simulation methods.** The wide range of implications of the expected findings are clearly dominated by the generalizability of the analysis method and its immanent synthesizing capabilities.



Multi-Layered Nested Mesh Processing